

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2003, 13:06:36 ; Search time 21 Seconds  
(without alignments)  
720.898 Million cell updates/sec

Title: US-09-831-426C-2  
Perfect score: 2008  
Sequence: 1 MDPFAVAVSVSLTIDAF.....ESPNCVEDKMLSTVAVLTIG 365

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID         | Description        |
|------------|--------|-------------|--------|------------|--------------------|
| 1          | 1993   | 99.3        | 423    | TF3A_HUMAN | 092664 homo sapien |
| 2          | 1055.5 | 52.6        | 366    | TF3A_XENLA | P03001 xenopus lae |
| 3          | 1047   | 52.1        | 335    | TF3A_RANPI | P34695 rana pipien |
| 4          | 1024   | 51.0        | 339    | TF3A_XENBO | P17842 xenopus bor |
| 5          | 1022   | 50.9        | 339    | TF3A_BUFAM | P34694 bufu americ |
| 6          | 816    | 40.6        | 322    | TF3A_ICTPU | P79797 ictalurus p |
| 7          | 556.5  | 27.7        | 365    | P43_XENLA  | P25456 xenopus lae |
| 8          | 553.5  | 27.6        | 365    | P43_XENBO  | P25066 xenopus bor |
| 9          | 417.5  | 20.8        | 803    | 2XDB_HUMAN | P98169 homo sapien |
| 10         | 410.5  | 20.4        | 616    | 2N93_HUMAN | P35789 homo sapien |
| 11         | 409    | 20.4        | 336    | 2G57_XENLA | P18729 xenopus lae |
| 12         | 409    | 20.4        | 799    | 2XDA_HUMAN | P98168 homo sapien |
| 13         | 406    | 20.2        | 698    | 2Z34_HUMAN | P14588 homo sapien |
| 14         | 404    | 20.1        | 821    | 2NA1_HUMAN | P51814 homo sapien |
| 15         | 403    | 20.1        | 574    | YF73_HUMAN | Q9P255 homo sapien |
| 16         | 399    | 19.9        | 1167   | 2Z08_HUMAN | Q43345 homo sapien |
| 17         | 398.5  | 19.8        | 626    | 2143_HUMAN | P52747 homo sapien |
| 18         | 398.5  | 19.8        | 751    | 2337_HUMAN | Q9Y3M9 homo sapien |
| 19         | 395.5  | 19.7        | 453    | 206_XENLA  | P18749 xenopus lae |
| 20         | 395    | 19.7        | 803    | 2Z26_HUMAN | Q9NYT6 homo sapien |
| 21         | 394.5  | 19.6        | 803    | 2NA3_HUMAN | P17038 homo sapien |
| 22         | 393    | 19.6        | 489    | 2ET1_HUMAN | Q9NQZ8 homo sapien |
| 23         | 393    | 19.6        | 689    | PRD9_HUMAN | Q9NQV7 homo sapien |
| 24         | 392.5  | 19.5        | 337    | 2N81_HUMAN | P51508 homo sapien |
| 25         | 391    | 19.5        | 738    | 2N84_HUMAN | P51523 homo sapien |
| 26         | 386.5  | 19.2        | 759    | 2Z87_MOUSE | Q9EQB9 homo sapien |
| 27         | 386.5  | 19.2        | 898    | 2Z07_XENLA | P18751 xenopus lae |
| 28         | 386    | 19.2        | 913    | 2Z28_HUMAN | Q9YJN3 homo sapien |
| 29         | 385    | 19.2        | 626    | 2189_HUMAN | Q75820 homo sapien |
| 30         | 383    | 19.1        | 538    | 2155_HUMAN | Q12901 homo sapien |
| 31         | 382.5  | 19.0        | 280    | 2G46_XENLA | P18722 xenopus lae |
| 32         | 382    | 19.0        | 645    | 2F93_MOUSE | Q61116 mus musculu |
| 33         | 381.5  | 19.0        | 535    | 2Z57_HUMAN | Q9Y291 homo sapien |

|    |       |      |     |   |            |                    |
|----|-------|------|-----|---|------------|--------------------|
| 34 | 381.5 | 19.0 | 734 | 1 | 2NA2_HUMAN | P28698 homo sapien |
| 35 | 381   | 19.0 | 711 | 1 | 2175_HUMAN | Q9Y473 homo sapien |
| 36 | 380.5 | 18.9 | 595 | 1 | 2N85_HUMAN | Q03923 homo sapien |
| 37 | 380.5 | 18.9 | 682 | 1 | 2NA5_HUMAN | Q02386 homo sapien |
| 38 | 379   | 18.9 | 683 | 1 | Y972_HUMAN | Q9Y2H8 homo sapien |
| 39 | 378   | 18.8 | 652 | 1 | Y798_HUMAN | Q94892 homo sapien |
| 40 | 377.5 | 18.8 | 595 | 1 | 2317_HUMAN | Q96P96 homo sapien |
| 41 | 377   | 18.8 | 604 | 1 | 2300_HUMAN | Q96P96 homo sapien |
| 42 | 376.5 | 18.8 | 446 | 1 | 2N70_HUMAN | Q9UC06 homo sapien |
| 43 | 376.5 | 18.8 | 472 | 1 | 2IK3_HUMAN | Q96P96 homo sapien |
| 44 | 376.5 | 18.8 | 570 | 1 | 2N76_HUMAN | P36508 homo sapien |
| 45 | 376.5 | 18.8 | 697 | 1 | HKR1_HUMAN | P10072 homo sapien |

## ALIGNMENTS

RESULT 1  
TF3A\_HUMAN STANDARD: PRT: 423 AA.  
AC Q92664: Q13097: Q12963:  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Transcription factor IIA (Factor A) (TFIIIA).  
GN GTF3A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Fetal brain;  
RX MEDLINE=95309028; PubMed=7789179;  
RA Arakawa H., Nagase H., Hayashi N., Ogawa M., Nagata M.,  
RT Fujimura T., Takahashi E., Shin S., Nakamura Y.;  
RT "Molecular cloning, characterization, and chromosomal mapping of a  
RT novel human gene (GTF3A) that is highly homologous to Xenopus  
RT transcription factor IIA.";  
RL Cytogenet. Cell Genet. 70:235-238(1995).  
RN [2]  
RP SEQUENCE OF 61-423 FROM N.A.  
RC TISSUE-Brain;  
RX MEDLINE=95347600; PubMed=7622052;  
RA Drew P.D., Nagle J.W., Canning R.D., Ozato K., Biddison W.E.,  
RT Becker K.G.;  
RT "Cloning and expression analysis of a human cDNA homologous to  
RT Xenopus TFIIIA.";  
RL Gene 159:215-218(1995).  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=94342241; PubMed=8063702;  
RA Moorefield B., Roeder R.G.;  
RT "Purification and characterization of human transcription factor  
RT IIA.";  
RL J. Biol. Chem. 269:20857-20865(1994).  
CC - FUNCTION: INTERACTS WITH THE INTERNAL CONTROL REGION (ICR) OF  
CC APPROXIMATELY 50 BASES WITHIN THE 5S RNA GENES, IS REQUIRED FOR  
CC CORRECT TRANSCRIPTION OF THESE GENES BY RNA POLYMERASE III. ALSO  
CC BINDS THE TRANSCRIBED 5S RNA'S. MAY INITIATE TRANSCRIPTION OF THE  
CC 5S RIBOSOMAL RNA GENE AND MAINTAIN THE STABILITY OF TRANSCRIPTION  
CC OF OTHER GENES.  
CC - SUBCELLULAR LOCATION: Nuclear.  
CC - TISSUE SPECIFICITY: UBICUITOUS.  
CC -  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: D32257; BAA06988.1; .  
 DR EMBL: U20272; AAA75623.1; .  
 DR EMBL: U14134; AAA21873.1; .  
 DR HSSP: P03001; 1TF3.  
 DR TRANSFAC: T04953; .  
 DR Genew: HGNC:4662; GTF3A.  
 DR MIM: 600860; .  
 DR InterPro: IPR000822; ZnF\_C2H2.  
 DR Pfam: PF00096; zf-C2H2; 9.  
 DR PRINTS: PR00048; ZINC\_FINGER.  
 DR SMART: SM00355; ZnF\_C2H2; 8.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 7.  
 DR PROSITE: PS50157; ZINC\_FINGER\_C2H2\_2; 8.  
 DR Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;  
 KM RNA-binding; Repeat; Nuclear protein.  
 FT DOMAIN 98 359 ZINC\_FINGERS.  
 FT ZN\_FING 98 122 C2H2-TYPE.  
 FT ZN\_FING 128 152 C2H2-TYPE.  
 FT ZN\_FING 158 183 C2H2-TYPE.  
 FT ZN\_FING 190 212 C2H2-TYPE (ATYPICAL).  
 FT ZN\_FING 220 244 C2H2-TYPE.  
 FT ZN\_FING 247 271 C2H2-TYPE.  
 FT ZN\_FING 275 297 C2H2-TYPE.  
 FT ZN\_FING 304 329 C2H2-TYPE.  
 FT ZN\_FING 335 359 C2H2-TYPE.  
 FT ZN\_FING 359 378 N -> H (IN REF. 2).  
 FT CONFLICT 214 214 K -> KL (IN REF. 2; AAA21873).  
 FT CONFLICT 221 221 MISSING (IN REF. 2; AAA21873).  
 FT CONFLICT 223 248 SLASHSGYI -> EFGISQWY (IN REF. 1).  
 FT CONFLICT 378 387  
 SO SEQUENCE 423 AA; 46847 MM; A627D064A43FB6F0 CRC64;

Query Match 99.3%; Score 1993; DB 1; Length 423;  
 Best Local Similarity 99.28; Pred. No. 8.5e-146;  
 Matches 362; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MDPAVVAESVSLTIADAFIAAGSSAPTPPRPALPRFICSPDCSANYSKAWKLDAN 60  
 :|||||  
 Db 59 LDPRAVVAESVSLTIADAFIAAGSSAPTPPRPALPRFICSPDCSANYSKAWKLDAN 118  
 :|||||  
 OY 61 LCKHTGERPFVCDYEGCGKAFIRDYHLSRHILHTGEGKPFVCAATGCCDCKFNTKSNLKKH 120  
 :|||||  
 Db 119 LCKHTGERPFVCDYEGCGKAFIRDYHLSRHILHTGEGKPFVCAANGCDCKFNTKSNLKKH 178  
 :|||||  
 OY 121 FERKHNQOKOYICSEFEDCKKTFKKHQQLKHOCQHTNEPLFKCTOEGCGKHFASSPKLK 180  
 :|||||  
 Db 179 FERKHNQOKOYICSEFEDCKKTFKKHQQLKHOCQHTNEPLFKCTOEGCGKHFASSPKLK 238  
 :|||||  
 OY 181 RHAKAHGVCOCGCSFVAKTWTETLKHVRETHKEEILCEVCARKTFKRDYLYKQHMKTHA 240  
 :|||||  
 Db 239 RHAKAHGVCOCGCSFVAKTWTETLKHVRETHKEEILCEVCARKTFKRDYLYKQHMKTHA 298  
 :|||||  
 OY 241 PERDYCRCPRGCGRTYTVFNLSHLSFHEESRPVCEHAGCGKTFAMKOSLTRHAVV 300  
 :|||||  
 Db 299 PERDYCRCPRGCGRTYTVFNLSHLSFHEESRPVCEHAGCGKTFAMKOSLTRHAVV 358  
 :|||||  
 OY 301 HDPOKKMKLKVKKSRKSLASHLSGYIPPKKQGGSLCQNGESPNCVEDKMLSTVA 360  
 :|||||  
 Db 359 HDPOKKMKLKVKKSRKSLASHLSGYIPPKKQGGSLCQNGESPNCVEDKMLSTVA 418  
 :|||||  
 OY 361 VLTIG 365  
 :|||||  
 Db 419 VLTIG 423

RESULT 2  
 TF3A\_XENLA STANDARD: PRT: 366 AA.

ID TF3A\_XENLA STANDARD: PRT: 366 AA.  
 AC P03001; Q91856;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Transcription factor IIA (Factor A) (TFIIIA) (S-TFIIIA/O-TFIIIA).  
 OS Xenopus laevis (African clawed frog).

CC Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 CC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE OF 23-366 FROM N.A.  
 RX MEDLINE=85074456; PubMed=6210149;  
 RA Ginsberg A.M., King B.O., Roeder R.G.;  
 RT "Xenopus 5S gene transcription factor, TFIIIA: characterization of a  
 RT cDNA clone and measurement of RNA levels throughout development.";   
 RL Cell 39:479-489(1984).  
 RN [2]  
 RP SEQUENCE OF 23-366 FROM N.A.  
 RX MEDLINE=86176722; PubMed=3754326;  
 RA Yun Tso J., van den Berg J., Korn L.J.;  
 RT "Structure of the gene for Xenopus transcription factor TFIIIA.";  
 RL Nucleic Acids Res. 14:2187-2201(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 112-137.  
 RX MEDLINE=86312888; PubMed=3755818;  
 RA Taylor W., Jackson I.J., Slegel N., Kumar A., Brown D.D.;  
 RT "The developmental expression of the gene for TFIIIA in Xenopus  
 RT laevis.";  
 RL Nucleic Acids Res. 14:6185-6195(1986).  
 RN [4]  
 RP SEQUENCE OF 1-35 FROM N.A.  
 RX MEDLINE=89306603; PubMed=2744458;  
 RA Scotto K.W., Kaulen H., Roeder R.G.;  
 RT "Positive and negative regulation of the gene for transcription  
 RT factor IIA in Xenopus laevis oocytes.";  
 RL Genes Dev. 3:651-662(1989).  
 RN [5]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RX MEDLINE=91071587; PubMed=2253860;  
 RA Kim S.R., Darby M.K., Joho K.E., Brown D.D.;  
 RT "The characterization of the TFIIIA synthesized in somatic cells of  
 RL Xenopus laevis.";  
 RN Genes Dev. 4:1602-1610(1990).  
 RN [6]  
 RP REPEATS ANALYSIS.  
 RX MEDLINE=85284956; PubMed=4040853;  
 RA Miller J., McLachlan A.D., Klug A.;  
 RT "Repetitive zinc-binding domains in the protein transcription factor  
 RT IIA from Xenopus oocytes.";  
 RL EMBO J. 4:1609-1614(1985).  
 RN [7]  
 RP REPEATS ANALYSIS.  
 RX MEDLINE=85231235; PubMed=4007166;  
 RA Brown R.S., Sander C., Argos P.;  
 RT "The primary structure of transcription factor TFIIIA has 12  
 RT consecutive repeats.";  
 RL FEBS Lett. 186:271-274(1985).  
 RN [8]  
 RP REPEATS ANALYSIS.  
 RA Boehm S., Drescher B.;  
 RT "Multiple internal repeats within the structure of the 5S RNA/DNA  
 RT binding transcription factor TF-IIIA from Xenopus laevis.";  
 RL Studia Biophys. 107:237-247(1985).  
 RN [9]  
 RP STRUCTURE BY NMR OF 10-101.  
 RX MEDLINE=97397344; PubMed=9253405;  
 RA Foster M.P., Wuttke D.S., Radhakrishnan I., Case D.A.,  
 RT Gottesfeld J.M., Wright P.E.;  
 RT "Domain packing and dynamics in the DNA complex of the N-terminal  
 RT zinc fingers of TFIIIA.";  
 RL Nat. Struct. Biol. 4:605-608(1997).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 32-210.  
 RX MEDLINE=98169475; PubMed=9501194;  
 RA Nolte R.T., Conlin R.M., Harrison S.C., Brown R.S.;  
 RT "Differing roles for zinc fingers in DNA recognition: structure of a  
 RT six-finger transcription factor IIA complex.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:2938-2943(1998).

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:48:16 / Search time 3381.18 Seconds

(Without alignments)  
10957.076 Million cell updates/sec

Title: US-09-831-426c-3

Perfect score: 1273

Sequence: 1 atgcgcagcagcgcgcgcga.....cagtactacccttgctaa 1273

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*\*  
1: gb\_ba: \*\*  
2: gb\_hcg: \*\*  
3: gb\_in: \*\*  
4: gb\_om: \*\*  
5: gb\_ov: \*\*  
6: gb\_pat: \*\*  
7: gb\_ph: \*\*  
8: gb\_pl: \*\*  
9: gb\_pr: \*\*  
10: gb\_ro: \*\*  
11: gb\_sts: \*\*  
12: gb\_sy: \*\*  
13: gb\_un: \*\*  
14: gb\_vl: \*\*  
15: em\_ba: \*\*  
16: em\_fun: \*\*  
17: em\_hum: \*\*  
18: em\_in: \*\*  
19: em\_mu: \*\*  
20: em\_om: \*\*  
21: em\_or: \*\*  
22: em\_ov: \*\*  
23: em\_pat: \*\*  
24: em\_ph: \*\*  
25: em\_pl: \*\*  
26: em\_ro: \*\*  
27: em\_sts: \*\*  
28: em\_un: \*\*  
29: em\_vl: \*\*  
30: em\_htg\_hum: \*\*  
31: em\_htg\_inv: \*\*  
32: em\_htg\_other: \*\*  
33: em\_htg\_mus: \*\*  
34: em\_htg\_pln: \*\*  
35: em\_htg\_rod: \*\*  
36: em\_htg\_mam: \*\*  
37: em\_htg\_vrt: \*\*  
38: em\_sy: \*\*  
39: em\_htgo\_hum: \*\*  
40: em\_htgo\_mus: \*\*  
41: em\_htgo\_other: \*\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID       | Description        |
|------------|--------|-------------|--------|-------------|--------------------|
| 1          | 1233.2 | 96.9        | 1381   | 9 HUMITIA   | D32257 Human GTF3A |
| 2          | 1231.6 | 96.7        | 1399   | 6 E10962    | E10962 CDNA's enco |
| 3          | 1088.8 | 85.5        | 1173   | 9 HSU20272  | U20272 Human DNA/R |
| 4          | 919.2  | 72.2        | 1098   | 9 HSU14134  | U14134 Human trans |
| 5          | 889.8  | 69.9        | 1920   | 9 AK057993  | AK057993 Homo sapi |
| 6          | 719.2  | 56.3        | 1298   | 10 AF391799 | AF391799 Mus muscu |
| 7          | 691    | 54.3        | 1071   | 10 AF391798 | AF391798 Rattus no |
| 8          | 560.6  | 44.0        | 996    | 10 BC032292 | BC032292 Mus muscu |
| 9          | 470.2  | 36.9        | 97979  | 9 AC004739  | AC004739 Homo sapi |
| 10         | 470.2  | 36.9        | 193126 | 9 AC006045  | AC006045 Homo sapi |
| 11         | 452.6  | 35.6        | 111722 | 9 AC074390  | AC074390 Homo sapi |
| 12         | 452.6  | 35.6        | 206622 | 2 AC016444  | AC016444 Homo sapi |
| 13         | 447    | 35.1        | 199517 | 9 AL353741  | AL353741 Human DNA |
| 14         | 418.4  | 32.9        | 439    | 6 AX396111  | AX396111 Sequence  |
| 15         | 416.8  | 32.7        | 439    | 6 AX396474  | AX396474 Sequence  |
| 16         | 385.4  | 30.3        | 163421 | 2 AC012664  | AC012664 Homo sapi |
| 17         | 385.4  | 30.3        | 197709 | 9 AC073641  | AC073641 Homo sapi |
| 18         | 376    | 29.5        | 166141 | 9 AL137059  | AL137059 Human DNA |
| 19         | 373.4  | 29.3        | 200885 | 2 AC012460  | AC012460 Homo sapi |
| 20         | 372.4  | 29.3        | 1693   | 2 AF265440  | AF265440 Homo sapi |
| 21         | 361.4  | 28.4        | 1518   | 5 XELTFITIA | XELTFITIA          |
| 22         | 355    | 27.9        | 1377   | 5 XBF3A     | XBF3A              |
| 23         | 350.8  | 27.6        | 1331   | 5 XELFTNAB  | XELFTNAB           |
| 24         | 340.6  | 26.8        | 1297   | 5 RPTRFITIA | RPTRFITIA          |
| 25         | 340.6  | 26.8        | 1305   | 5 RANFTITIA | RANFTITIA          |
| 26         | 319.2  | 25.1        | 123905 | 9 AC116311  | AC116311 Homo sapi |
| 27         | 319.2  | 25.1        | 157165 | 2 AC104112  | AC104112 Homo sapi |
| 28         | 319.2  | 25.1        | 171007 | 2 AC021381  | AC021381 Homo sapi |
| 29         | 319.2  | 25.1        | 171705 | 2 AC113377  | AC113377 Homo sapi |
| 30         | 284.6  | 22.4        | 1314   | 5 BATFITIA  | BATFITIA           |
| 31         | 278.6  | 21.9        | 67190  | 2 AC117566  | AC117566 Mus muscu |
| 32         | 234    | 18.4        | 1129   | 6 AX396395  | AX396395 Sequence  |
| 33         | 221.6  | 17.4        | 1129   | 5 IPOOTF3A  | IPOOTF3A           |
| 34         | 202    | 15.9        | 162774 | 2 AC121381  | AC121381 Rattus no |
| 35         | 202    | 15.9        | 182326 | 2 AC125693  | AC125693 Rattus no |
| 36         | 197.6  | 15.5        | 112902 | 9 AC005230  | AC005230 Homo sapi |
| 37         | 177    | 13.9        | 272    | 6 AX322436  | AX322436 Sequence  |
| 38         | 163    | 12.8        | 167342 | 2 AC124828  | AC124828 Mus muscu |
| 39         | 163    | 12.8        | 217225 | 10 AL513345 | AL513345 Mouse DNA |
| 40         | 152.4  | 12.0        | 660    | 6 AX202580  | AX202580 Sequence  |
| 41         | 152.4  | 12.0        | 947    | 6 AX053174  | AX053174 Sequence  |
| 42         | 152.4  | 12.0        | 947    | 6 AX059770  | AX059770 Sequence  |
| 43         | 152.4  | 12.0        | 947    | 6 AX201763  | AX201763 Sequence  |
| 44         | 152.4  | 12.0        | 947    | 6 AX201936  | AX201936 Sequence  |
| 45         | 152.4  | 12.0        | 995    | 6 AX053173  | AX053173 Sequence  |

## ALIGNMENTS

RESULT 1  
HUMITIA  
LOCUS  
DEFINITION Human GTF3A mRNA for xenopus transcription factor IIAA homologue,  
complete cds.  
ACCESSION D32257.1 GI:1000446  
VERSION D32257.1  
KEYWORDS GTF3A; Xenopus transcription factor IIAA homologue.  
SOURCE Homo sapiens CDNA to mRNA, clone\_11b:library of T.Fujiwara, S.Shin  
and Y.Nakamura clone:39H11.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1381)  
Arikawa,H., Nagase,H., Hayashi,N., Ogawa,M., Nagata,M.,

| TITLE  | JOURNAL                | REFERENCE                |
|--|------------------------|--------------------------|
| Fujiwara, T., Takahashi, E., Shio, S. and Nakamura, Y. Molecular cloning, characterization, and chromosomal mapping of a novel human gene (GTF3A) that is highly homologous to Xenopus transcription factor IIIA | Cytogenet. Cell Genet. | 70 (3-4), 235-238 (1995) |
|  |                        | 95309028                 |
|  |                        | 1 (bases 1 to 1381)      |

Submitted: 22-Jul-1994) Yusuke Nakamura, Cancer Institute,  
Department of Biochemistry, 1-37-1 Kami-Kobukuro, Toshima-ku  
Tokyo 170, Japan (E-mail: nakamura@ganv.tci.ac.jp,  
Tel: 03-3918-0111 (ex. 4501), Fax: 03-3918-0342)  
Received: 20-Jul-1994

**source**

```

gene      /clone_lib="Library of T.Fujiwara, S.Shin and Y.Nakamura
          1. .1381
          /gene="CTF3A"
CDS       20. .1291
          /start="start"

```

[illegible]

|                       |              |                     |               |              |
|-----------------------|--------------|---------------------|---------------|--------------|
| Query Match           | 96.98%       | Score 1233.2;       | DB 9;         | Length 1361; |
| Best Local Similarity | 99.50%       | Pred. NO. 4.1e-259; |               |              |
| Matches 1268;         | Conservative | 0;                  | Mismatches 3; | Indels 3;    |
|                       |              |                     | Gaps          | 3            |

|    |     |   |     |
|----|-----|---|-----|
| QY | 1   | ATGCGCAGCAGCGCCCGCCGACCGCGGGGCGGTGCCTGTGACCGCGCGCTCCCGCAAGT     | 60  |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| Db | 20  | ATGCGCAGCAGCGCGCCCGACCGCGGGCGGTGCCTGTGACCGCGCGCTCCCGCAAGT       | 79  |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| QY | 61  | GTCGCGCGGTGCGGGCAAGGTTTCAGCAGAGCGCTGTGGCCGCGCGCGCTTCCCGGC       | 120 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| Db | 80  | GTCGCGCGGTGCGGGCAAGGTTTCAGCAGAGCGCTGTGGCCGCGCGCGCTTCCCGGC       | 138 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| QY | 121 | ACGTGTCTCGGGCAGCTGGCAGCGCGCCCTGGCCCTTGGGCTTGGAGCGCGCGCGCCCTGGA  | 180 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| Db | 139 | ACGTGTCTCGGGCAGCTGGCAGCGCGCCCTGGCCCTTGGGAGCGCGCGCGCCCTGGA       | 198 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| QY | 181 | TCCGCGCGCGGTGTGCGCCGAGTCGCGGTGTCTGTGACCATGCGCGGACCGCTTCATTCG    | 240 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| Db | 199 | TCCGCGCGCGGTGTGCGCCGAGTCGCGGTGTCTGTGACCATGCGCGGACCGCTTCATTCG    | 258 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| QY | 241 | AGCCGGCGAGAGCTCAGCTCCGACCCCGCGCGCGCGCTTCCAGAGAGTTCACTGTG        | 300 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| Db | 259 | AGCCGGCGAGAGCTCAGCTCCGACCCCGCGCGCGCGCTTCCAGAGAGTTCACTGTG        | 318 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| QY | 301 | CTCCTTCCCTGACTGCAAGCGCCCAATTACAGCAAAAGCCTGGAAGCTTGACCGGCACCTGTG | 360 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| Db | 319 | CTCCTTCCCTGACTGCAAGCGCCCAATTACAGCAAAAGCCTGGAAGCTTGACCGGCACCTGTG | 378 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| QY | 361 | CAAGCAGACGGGGGAGAGACCAATTGTTGTGACTATGAAGGCTGTGGCAAGCCCTTCAT     | 420 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| Db | 379 | CAAGCAGACGGGGGAGAGACCAATTGTTGTGACTATGAAGGCTGTGGCAAGCCCTTCAT     | 438 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| QY | 421 | CAGGAGCTAACCATCTGAGCGCGCCACATTCTGACTCACACAGGAGAAAAAGCGTTGTGTG   | 480 |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
|    |     |   |     |
| Db | 439 | CAGGAGCTAACCATCTGAGCGCGCCACATTCTGACTCACACAGGAGAAAAAGCGTTGTGTG   | 498 |

|    |      |  |      |
|----|------|--|------|
| OY | 481  | TGCAGCCACTGCCTGTGATCCAAAAATTCACACAAAAATCAAACCTGTGAAGAAACATTTTGA  | 540  |
|    |      |  |      |
| DB | 499  | TGCAGCCAAATGGCTGTGATCCAAAAATTCACACAAAAATCAAACCTGTGAAGAAACATTTTGA | 558  |
| OY | 541  | ACGCAAAACATGAAAAATCAACAAAAAACATATATATGCACTTTTGAAGACTGTAAAGAAC    | 600  |
|    |      |  |      |
| DB | 559  | ACGCAAAACATGAAAAATCAACAAAAAACATATATATGCACTTTTGAAGACTGTAAAGAAC    | 618  |
| OY | 601  | CTTTAAGAAACATCAGCAGCTGAAAAATCCATCAGTGGCCAGCATACCAATGAACCTCTATT   | 660  |
|    |      |  |      |
| DB | 619  | CTTTAAGAAACATCAGCAGCTGAAAAATCCATCAGTGGCCAGCAATACCAATGAACCTCTATT  | 678  |
| OY | 661  | CAAGTGTACCCAGGAAGGATGTGGGAAACACTTTGCATCACCCAGCAAGCTGAAACGACA     | 720  |
|    |      |  |      |
| DB | 679  | CAAGTGTACCCAGGAAGGATGTGGGAAACACTTTGCATCACCCAGCAAGCTGAAACGACA     | 738  |
| OY | 721  | TGCCAAGGCCACAGAGGGCTATGTATGTCAAAAAAGATCTTCCTTTGTGGCAAAAAACATG    | 780  |
|    |      |  |      |
| DB | 739  | TGCCAAGGCCACAGAGGGCTATGTATGTCAAAAAAGATCTTCCTTTGTGGCAAAAAACATG    | 798  |
| OY | 781  | GACGGAACCTCTGAAACATGTGAGAGAAACCCATAAAGAGAAATACTATGTCAAGTATG      | 840  |
|    |      |  |      |
| DB | 799  | GACGGAACCTCTGAAACATGTGAGAGAAACCCATAAAGAGAAATACTATGTCAAGTATG      | 858  |
| OY | 841  | CCGGAATAACATTTTAAACGCAAAAGATTACCTTAAGCAACACATGAAAACTCATGCCCCAGA  | 900  |
|    |      |  |      |
| DB | 859  | CCGGAATAACATTTTAAACGCAAAAGATTACCTTAAGCAACACATGAAAACTCATGCCCCAGA  | 918  |
| OY | 901  | AAGGATGTATGTGCGCTGTGCCAGAGAGCGCTGTGGAAGACCTATACTACTGTGTTAA       | 960  |
|    |      |  |      |
| DB | 919  | AAGGATGTATGTGCGCTGTGCCAGAGAGCGCTGTGGAAGACCTATACTACTGTGTTAA       | 978  |
| OY | 961  | TCTCCAAGCCATATCCTCTCCTTCATGAGGAAAGCCGCCCTTTGTGTGAACATGC          | 1020 |
|    |      |  |      |
| DB | 979  | TCTCCAAGCCATATCCTCTCCTTCATGAGGAAAGCCGCCCTTTGTGTGAACATGC          | 1030 |
| OY | 1021 | TGGCTGTGGCAAAAACATTTGCAATGAAACAAAGTCTCAGTAAGGCATGCTGTTGTACATGA   | 1080 |
|    |      |  |      |
| DB | 1039 | TGGCTGTGGCAAAAACATTTGCAATGAAACAAAGTCTCAGTAAGGCATGCTGTTGTACATGA   | 1090 |
| OY | 1081 | TCCTGACAAAGAAATGAAAGCTCAAAAGTCAAAAATCTCGTGAAAAAC -GAGTTTGG       | 1130 |
|    |      |  |      |
| DB | 1099 | TCCTGACAAAGAAATGAAAGCTCAAAAGTCAAAAATCTCGTGAAAAACGCGAGATTGG       | 1150 |
| OY | 1140 | CCTCTCATCTCAGTGGATATATCCCTGCCAAAAAGGAACAGGGCAAGCCTTATCTTTGT      | 1190 |
|    |      |  |      |
| DB | 1159 | CCTCTCATCTCAGTGGATATAT - CCTGCCAAAAAGGAACAGGGCAAGCCTTATCTTTGT    | 1210 |
| OY | 1200 | GTCAAAACGGAGAGATCACCCCACTGTGTGGAAGACAAAGATGCTCTGACAGTTGCAGTAC    | 1250 |
|    |      |  |      |
| DB | 1218 | GTCAAAACGGAGAGATCACCCCACTGTGTGGAAGACAAAGATGCTCTGACAGTTGCAGTAC    | 1270 |
| OY | 1260 | TTACCCCTTGGCTAA  | 1273 |
|    |      |  |      |
| DB | 1278 | TTACCCCTTGGCTAA  | 1291 |

|            |   |             |     |        |
|------------|---|-------------|-----|--------|
| RESULT 2   |   |             |     |        |
| E10962     |   |             |     |        |
| LOCUS      | E10962  | 1399 bp     | RNA | linear |
| DEFINITION | CDNA's encoding htfIIIA protein.                                  |             |     |        |
| ACCESSION  | E10962  |             |     |        |
| VERSION    | E10962.1  | GI:22028832 |     |        |
| KEYWORDS   | JP 1996070870-A/1.  |             |     |        |
| SOURCE     | Homo sapiens.   |             |     |        |
| ORGANISM   | Homo sapiens  |             |     |        |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |             |     |        |
|            | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.         |             |     |        |
| REFERENCE  | 1 (bases 1 to 1399)   |             |     |        |
| AUTHORS    | Fujiwara,T., Takeda,S., Shimada,Y., Ozaki,K. and Shln,T.          |             |     |        |
| TITLE      | HUMAN TRANSCRIPTIONAL FACTOR IIA GENE                             |             |     |        |
| JOURNAL    | Patent: JP 1996070870-A 1 19-MAR-1996;                            |             |     |        |

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:48:16 ; Search time 3221.82 Seconds

(without alignments)  
10957.076 million cell updates/sec

Title: US-09-831-426C-4

Sequence: 1 gtgccgagccgcgcgaag.....cagtacttacccttgctaa 1213

Scoring table: IDENTITY\_NUC

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

```

1:  gb_ba:*
2:  gb_hcg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_ln:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pln:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcg_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID           | Description         |
|------------|--------|-------------|--------|--------------|---------------------|
| 1          | 1171.6 | 56.6        | 1381   | 9 HUMTITA    | D32257 Human GTF3A  |
| 2          | 1170   | 96.5        | 1399   | 6 E10962     | E10962 CDNA's enclo |
| 3          | 1088.8 | 89.8        | 1173   | 9 HSU20272   | U20272 Human DNA/R  |
| 4          | 919.2  | 75.8        | 1098   | 9 AKS014134  | U14134 Human trans  |
| 5          | 889.8  | 73.4        | 1920   | 9 HSU57993   | AK057993 Homo sapi  |
| 6          | 719.2  | 59.3        | 1298   | 10 AF391799  | AF391799 Mus muscu  |
| 7          | 691    | 57.0        | 1071   | 10 AF391798  | AF391798 Rattus no  |
| 8          | 560.6  | 46.2        | 996    | 10 BC032292  | BC032292 Mus muscu  |
| 9          | 470.2  | 38.8        | 97979  | 9 AC004739   | AC004739 Homo sapi  |
| 10         | 470.2  | 38.8        | 193126 | 9 AC006045   | AC006045 Homo sapi  |
| 11         | 452.6  | 37.3        | 111722 | 9 AC074390   | AC074390 Homo sapi  |
| 12         | 452.6  | 37.3        | 206622 | 2 AC016444   | AC016444 Homo sapi  |
| 13         | 447    | 36.9        | 199519 | 9 AL353741   | AL353741 Human DNA  |
| 14         | 418.4  | 34.5        | 439    | 6 AX396111   | AX396111 Sequence   |
| 15         | 416.8  | 34.4        | 439    | 6 AX396474   | AX396474 Sequence   |
| 16         | 385.4  | 31.8        | 163421 | 2 AC012664   | AC012664 Homo sapi  |
| 17         | 385.4  | 31.8        | 197709 | 9 AC073641   | AC073641 Homo sapi  |
| 18         | 373.4  | 30.8        | 200885 | 2 AC012460   | AC012460 Homo sapi  |
| 19         | 372.4  | 30.7        | 1693   | 9 AF265440   | AF265440 Homo sapi  |
| 20         | 361.4  | 29.8        | 1518   | 5 XELTPTITA  | K02938 X.laevis 5S  |
| 21         | 355    | 29.3        | 1377   | 5 XBTFF3A    | X17695 Xenopus bor  |
| 22         | 350.8  | 28.9        | 1331   | 5 XELFINAB   | M32472 X.borealis   |
| 23         | 340.6  | 28.1        | 1297   | 5 RPTRETTITA | X58369 R.pilpiens m |
| 24         | 340.6  | 28.1        | 1305   | 5 RANPTTITA  | M85211 Rana pipien  |
| 25         | 319.2  | 26.3        | 123905 | 9 AC116311   | AC116311 Homo sapi  |
| 26         | 319.2  | 26.3        | 157165 | 2 AC104112   | AC104112 Homo sapi  |
| 27         | 319.2  | 26.3        | 171007 | 2 AC021381   | AC021381 Homo sapi  |
| 28         | 319.2  | 26.3        | 171705 | 2 AC113377   | AC113377 Homo sapi  |
| 29         | 314.4  | 25.9        | 166141 | 5 AL137059   | AL137059 Human DNA  |
| 30         | 284.6  | 23.5        | 1314   | 5 BATFTTITA  | X58366 B.americanu  |
| 31         | 278.6  | 23.0        | 67190  | 2 AC117566   | AC117566 Mus muscu  |
| 32         | 234    | 19.3        | 234    | 6 AX396395   | AX396395 Sequence   |
| 33         | 221.6  | 18.3        | 1129   | 5 IPOQFF3A   | Z68499 I.punctatus  |
| 34         | 202    | 16.7        | 162774 | 2 AC121381   | AC121381 Rattus no  |
| 35         | 202    | 16.7        | 182326 | 2 AC125693   | AC125693 Rattus no  |
| 36         | 197.6  | 16.3        | 112902 | 6 AC005230   | AC005230 Homo sapi  |
| 37         | 177    | 14.6        | 272    | 6 AX322436   | AX322436 Sequence   |
| 38         | 163    | 13.4        | 167342 | 2 AC124828   | AC124828 Mus muscu  |
| 39         | 163    | 13.4        | 217225 | 10 AL513345  | AL513345 Mouse DNA  |
| 40         | 152.4  | 12.6        | 660    | 6 AX202580   | AX202580 Sequence   |
| 41         | 152.4  | 12.6        | 947    | 6 AX053174   | AX053174 Sequence   |
| 42         | 152.4  | 12.6        | 947    | 6 AX059770   | AX059770 Sequence   |
| 43         | 152.4  | 12.6        | 947    | 6 AX201763   | AX201763 Sequence   |
| 44         | 152.4  | 12.6        | 947    | 6 AX201936   | AX201936 Sequence   |
| 45         | 152.4  | 12.6        | 995    | 6 AX053173   | AX053173 Sequence   |

## ALIGNMENTS

| RESULT 1   |   |
|------------|---|
| LOCUS      | HMU111A   |
| DEFINITION | Human GTF3A mRNA for Xenopus transcription factor IIA homologue, complete cds.  |
| ACCESSION  | D33257  |
| VERSION    | GI:1000446  |
| KEYWORDS   | GT3A; Xenopus transcription factor IIA homologue.   |
| SOURCE     | Homo sapiens cDNA to mRNA, clone_11b:library of T.Fujiwara, S.Shin and Y.Nakamura clone:39H11.                              |
| ORGANISM   | Homo sapiens  |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| AUTHORS    | 1 (bases 1 to 1381)<br>Atakawa,H., Nagase,H., Hayashi,N., Ogawa,M., Nagata,M.,  |

```

TITLE      Fujiwara,T., Takahashi,E., Shin,S. and Nakamura,Y.
            Molecular cloning, characterization, and chromosomal mapping of a
            novel human gene (GFP3A) that is highly homologous to Xenopus
            transcription factor IIA
JOURNAL     Cytogenet. Cell Genet. 70 (3-4), 235-238 (1995)
MEDLINE     95309028
REFERENCE   2 (bases 1 to 1381)
AUTHORS     Nakamura,Y.
TITLE       Direct Submission
JOURNAL     Submitted (22-JUL-1994) Yusuke Nakamura, Cancer Institute,
            Department of Biochemistry; 1-37-1 Kami-Ikebukuro, Toshima-ku,
            Tokyo 170, Japan (E-mail:nakamura@ganvxi.jicr.or.jp,
            Tel:03-3918-0111(ex.4501), Fax:03-3918-0342)
FEATURES
source
1..1381
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="39H1"
/clone_1db="Library of T.Fujiwara, S.Shin and Y.Nakamura"
1..1381
/gene="GFP3A"
20..1291
/gene="GFP3A"
/codon_start=1
/product="Xenopus transcription factor IIA homologue"
/protein_id="BAA06988.1"
/db_xref="GI:1616942"
/translation="MRSSGADAGRCILVTARAGSVPARSGSAGSRGARFPARYSA
RGSARPGLGAGALDPPRAVVAESVSLTIADAFIAAGESSAPTPRALPRFICSF
PDCSANSKAMKLDAILCKHNTGERPFVCDYEGCGKAFIRDYLSRHLTHTGEKPEVC
ANGCDOKENTKSNLKKFERKHNQKOYICSPEDCKKTFKKHQQLIHOCQTNPEP
LFKCTOECCKKHPASPSKLRNAKHAHEGYVCOXGCSFVAKTWTLSLHVRETHKEIL
CEVCRKTFPRKDYLRKOHMKTHADERDVCRCREGGRTYTFVNIQSHILSPHEBRP
FVCEHAGCGCTFAMKOSLTRHAYVHDPDKMKMLKVKRSREKREFGSSQWITPPKRRK
OGGSLCQNGESPNCVEDKMLSTVAVLTTG"
BASE COUNT 386 a 350 c 354 g 291 t
ORIGIN
Query Match      96.6%  Score 1171.6;  DB 9;  Length 1381;
Best Local Similarity 99.4%  Pred. No. 4.6e-253;
Matches 1207;  Conservative 0;  Mismatches 4;  Indels 3;  Gaps 3;
QY      1  GTGCCCGCGCCGCGGGAAGGTTACAGCAGGAGGCGGTGGCGCGCGCGCGCTCCCGGC 60
            |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB      80  GTGCCCGCGCTGCGCGGGAAGGTTACAGCAGGAGGCGGTGGCGCGCGCGCGCTCCCGGC 138
QY      61  AGCTGTCTCGGCACAGTGGCAGCGCGCGCTGGCCCTGGCGCTTGAGAGCGCGCGCGCGCTGGA 120
            |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB     139  AGGTGTCTCGGCACGTGGCAGCGCGCGCTGGCCCTGGCGCTTGAGAGCGCGCGCGCGCTGGA 198
QY     121  TCCGCGCGCGCGTGTGCGCGGAGTGGTGTGTCCTTGACCATGCGCGCGCTTCATTTGC 180
            |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB     199  TCCGCGCGCGCGTGTGCGCGGAGTGGTGTGTCCTTGACCATGCGCGCGCTTCATTTGC 258
QY     181  AGCCGGCGAGAGCTCAGCTCCGACCCCGCGCGCGCGCGCGCTTCCAGGAGGTTCAATCTG 240
            |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB     259  AGCCGGCGAGAGCTCAGCTCCGACCCCGCGCGCGCGCGCGCTTCCAGGAGGTTCAATCTG 318
QY     241  CTCCTTCCCTGACTCGAGCGCCCAATTACAGCAAGCCTGGAAGCTTGAGCGCGCACCTGTG 300
            |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB     319  CTCCTTCCCTGACTCGAGCGCCCAATTACAGCAAGCCTGGAAGCTTGAGCGCGCACCTGTG 378
QY     301  CAAGCAGACGGGGGAGACACCATTTGTTGTGACTATGAAGGGTGTGGCAAGGCCCTTCAT 360
            |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB     379  CAAGCAGACGGGGGAGAGACCATTTGTTGTGACTATGAAGGGTGTGGCAAGGCCCTTCAT 438
QY     361  CAGGACTACCATCTGAGCGCGCACATTTCTGACTACACAGAGAGAAAAGCCGTTGTTTG 420
            |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB     439  CAGGACTACCATCTGAGCGCGCACATTTCTGACTACACAGAGAGAAAAGCCGTTGTTTG 498
QY     421  TGCAGGCACTGGCTGTGATCAAAAATTCAACACAAAATCAAACTTGAAGAAACATTTTGA 480
            |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB     499  TGCAGGCACTGGCTGTGATCAAAAATTCAACACAAAATCAAACTTGAAGAAACATTTTGA 558

```

|            |  |   |                            |
|------------|--|---|----------------------------|
| OY         | 481  | ACGCAAAACATGAATAATCAACAAGAACATATATATATGCAGCTTTTGAAGACTGTAGAAGAC | 540                        |
| Db         | 559  | ACGCAAAACATGAATAATCAACAAGAACATATATATATGCAGCTTTTGAAGACTGTAGAAGAC | 618                        |
| OY         | 541  | CTTTAAGAAACATCAGCACGCTGAAAATCCATCAGTGCGCAGCATACCAATGAACCTCTATT  | 600                        |
| Db         | 619  | CTTTAAGAAACATCAGCACGCTGAAAATCCATCAGTGCGCAGCATACCAATGAACCTCTATT  | 678                        |
| OY         | 501  | CAAGTGTACCAGGAGGATGTGGGAAACACTTTGCATCAACCCAGCAGCTGAACGACA       | 660                        |
| Db         | 679  | CAAGTGTACCAGGAGGATGTGGGAAACACTTTGCATCAACCCAGCAGCTGAACGACA       | 738                        |
| OY         | 661  | TGCCAAGCCCCACGAGGGCTATGTATGTCAAAAAGGATGTCCTTTGTGGCAAAAAATG      | 720                        |
| Db         | 739  | TGCCAAGCCCCACGAGGGCTATGTATGTCAAAAAGGATGTCCTTTGTGGCAAAAAATG      | 798                        |
| OY         | 721  | GACGGAACCTCTGAAACATGTGAGAGAAACCCATTAAGAGGAATCTATGTGAAGTATG      | 780                        |
| Db         | 799  | GACGGAACCTCTGAAACATGTGAGAGAAACCCATTAAGAGGAATCTATGTGAAGTATG      | 858                        |
| OY         | 781  | CCGCAAAACATTTAAACGCAAGATTACCTTAAACAACATGAANAATCATGCCCCAGA       | 840                        |
| Db         | 859  | CCGCAAAACATTTAAACGCAAGATTACCTTAAACAACATGAANAATCATGCCCCAGA       | 918                        |
| OY         | 841  | AAGGATGTATGTGCGCTGTCCAAGAGAGGCTGTGGAAGAACCTATACTACTGTGTTAA      | 900                        |
| Db         | 919  | AAGGATGTATGTGCGCTGTCCAAGAGAGGCTGTGGAAGAACCTATACTACTGTGTTAA      | 978                        |
| OY         | 901  | TCCTCAAAGCCATATCCTCTCTCCATGAGGAAAGCCGCCCTTTGTGTGAACATGC         | 960                        |
| Db         | 979  | TCCTCAAAGCCATATCCTCTCTCCATGAGGAAAGCCGCCCTTTGTGTGAACATGC         | 1038                       |
| OY         | 961  | TGCGTGTGGCAAAACATTTGCAATGAAACAAGTCTCCTAGGCACTGCTGTGTACATGA      | 1020                       |
| Db         | 1039   | TGCGTGTGGCAAAACATTTGCAATGAAACAAGTCTCCTAGGCACTGCTGTGTACATGA      | 1098                       |
| OY         | 1021   | TCCTGACAAAGAAATGAAGCTCAAAAGTCAAAAAATCTCGTGA AAAAC - GAGTGTGG    | 1079                       |
| Db         | 1099   | TCCTGACAAAGAAATGAAGCTCAAAAGTCAAAAAATCTCGTGA AAAAC GAGTGTGG      | 1158                       |
| OY         | 1080   | CCTCTCATCTCAGTGGATATATCCCTCCCAAAAGGAAACAGGGCAGGCTTATCTTTGT      | 1139                       |
| Db         | 1159   | CCTCTCATCTCAGTGGATATAT - CCTCCCAAAAGGAAACAGGGCAGGCTTATCTTTGT    | 1217                       |
| OY         | 1140   | GTCAAAACGAGAGCTCACCCAACTGTGTGGAAGACAAGATGCTCTCGACAGTTGCAGTAC    | 1199                       |
| Db         | 1218   | GTCAAAACGAGAGCTCACCCAACTGTGTGGAAGACAAGATGCTCTCGACAGTTGCAGTAC    | 1277                       |
| OY         | 1200   | TTACCCTTGCTAA 1213  |                            |
| Db         | 1278   | TTACCCTTGCTAA 1291  |                            |
| RESULT 2   | EI0962   |   |                            |
| LOCUS      | EI0962   | 1399 bp   | RNA linear PAT 29-SEP-1997 |
| DEFINITION | CDNA's encoding htfIIIA protein.                                   |   |                            |
| ACCESSION  | EI0962   |   |                            |
| VERSION    | EI0962.1 GI:22028822   |   |                            |
| KEYWORDS   | JP 1996070870-A/1.   |   |                            |
| SOURCE     | Homo sapiens.  |   |                            |
| ORGANISM   | Homo sapiens   |   |                            |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; |   |                            |
| AUTHORS    | Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.           |   |                            |
| TITLE      | Fujiwara,T., Takeda,S., Shimada,Y., Ozaki,K. and Shin,T.           |   |                            |
| JOURNAL    | HUMAN TRANSCRIPTOMIAL FACTOR IIA GENE                              |   |                            |
| COMMENT    | Patent: JP 1996070870-A 1 19-MAR-1996;                             |   |                            |
|            | OTSUKA PHARMACEUT CO LTD   |   |                            |
|            | OS Homo sapiens (human)  |   |                            |
|            | PN JP 1996070870-A/1   |   |                            |
|            | PD 19-MAR-1996   |   |                            |

DT 07-JUL-1996 (first entry)  
 XX  
 DE Transcription factor-IIIA.  
 XX  
 KM Human; transcription factor-IIIA; hTFIIIA; DNA binding protein;  
 KM ribosome; zinc finger; diagnostic; probe; transcription control;  
 KM antitumour; cancer; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP704526-A1.  
 XX  
 PD 03-APR-1996.  
 XX  
 PF 05-SEP-1995; 95SEP-0113908.  
 XX  
 PR 05-SEP-1994; 94JP-0211022.  
 XX  
 PA (SAKA ) OTSUKA PHARM CO LTD.  
 XX  
 PI Fujiwara T, Ozaki K, Shimada Y, Shin S, Takeda S;  
 XX  
 DR WPI; 1996-173033/18.  
 XX  
 DR N-PSDB; AAT14037, AAT14038.  
 XX  
 PT Human Transcription Factor III A gene - useful in regulation of  
 PT transcription and for diagnosis and treatment of e.g. cancer related  
 PT diseases  
 XX  
 PS Claim 1: Page 8-10; 17pp; English.  
 XX  
 CC The sequence represents human transcription factor-IIIA (hTFIIIA),  
 CC a DNA binding protein which is necessary for the initiation of 5S  
 CC rRNA gene transcription, binding to an internal control region of  
 CC the 5S gene. The protein contains 9 zinc finger domains, which are  
 CC homologous to the C2H2 finger domains of Xenopus TFIIIA, except for  
 CC the 6th finger domain, which has only 3 amino acid residues between  
 CC 2 cysteine residues, instead of 5 amino acid residues for Xenopus  
 CC TFIIIA. The protein (optionally in recombinant form) and encoding  
 CC gene may be used in diagnosis, identification or therapy of  
 CC hereditary diseases such as cancer, or other diseases resulting from  
 CC abnormal transcriptional control, and to analyse the mechanisms  
 CC involved in their activity.  
 XX  
 SQ Sequence 423 AA:  
 Query Match 96.3%; Score 1933; DB 17; Length 423;  
 Best Local Similarity 96.2%; Pred. NO. 5.5e-153;  
 Matches 351; Conservative 3; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 MDPRAVVAESVSLTIADAFIAGESSAPTPPRPALPRRFTCSFPDCSANTSKAWKLDAA 60  
 Db :|||||  
 Db 59 LDPRAVVAESVSLTIADAFIAGESSAPTPPRPALPRRFTCSFPDCSANTSKAWKLDAA 118  
 QY 61 LCKHTGERPVCYDEGCGKAFIRDYHLSRHILHTGGEKPFVCAATGCDQKFNKSNLKKH 120  
 Db :|||||  
 Db 119 LCKHTGERPVCYDEGCGKAFIRDYHLSRHILHTGGEKPFVCAANGCDQKFNKSNLKKH 178  
 QY 121 FERKHNQOKOYICSFEDCKTKFKKHQOLKHOCQHTNEPLFKCTOEGCGKHFASPSK 180  
 Db :|||||  
 Db 179 FERKHNQOKOYICSFEDCKTKFKKHQOKIHOCQHTNEPLFKCTOEGCGKHFASPSK 238  
 QY 181 RHAKAHEGYVCOGCSFVAKTWTELLKHVRETHKEILCEVGRKTKFKRKDYLKQHMKTHA 240  
 Db :|||||  
 Db 239 RHAKAHEGYVCOGCSFVAKTWTELLKHVRETHKEILCEVGRKTKFKRKDYLKQHMKTHA 298  
 QY 241 PERDVCRCPRREGCGRTYTVFNLSHLSFHEESRPVCEHAGCGKTFAMKQSLTRHAAV 300  
 Db :|||||  
 Db 299 PERDVCRCPRREGCGRTYTVFNLSHLSFHEESRPVCEHAGCGKTFAMKQSLTRHAAV 358  
 QY 301 HDPDKKKMKTLYKKSREKRLASHLSGVIIPRKQOGGLSLONGESPNCVEDKMLSTVA 360  
 Db :|||||  
 Db 359 HDPDKKKMKTLYKKSREKRLASHLSGVIIPRKQOGGLSLONGESPNCVEDKMLSTVA 418

QY 361 VLTLG 365  
 Db 419 VLTLG 423  
 RESULT 4  
 ABP41846  
 ID ABP41846 standard; protein; 184 AA.  
 XX  
 AC ABP41846;  
 XX  
 DT 22-AUG-2002 (first entry)  
 XX  
 DE Human ovarian antigen HUSYA18, SEQ ID NO:2978.  
 XX  
 KM Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KM ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KM infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KM PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KM inflammatory condition; immune disorder; blood disorder;  
 KM cardiovascular disorder; respiratory disorder; neurological disorder;  
 KM gastrointestinal disorder; urinary system disorder; drug screening;  
 KM gene therapy; chromosome mapping; forensic analysis;  
 KM antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KM antiinflammatory; gynaecological; reproductive; chromosome 13q12.3-13.1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200677-A1.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 07-JUN-2001; 2001WO-US18569.  
 XX  
 PR 07-JUN-2000; 2000US-209467P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Birse CE, Rosen CA;  
 XX  
 DR WPI; 2002-147878/19.  
 DR N-PSDB; ABQ54923.  
 XX  
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
 PT ovarian cancer), immune disorders, cardiovascular disorders and  
 PT neurological diseases .  
 XX  
 PS Claim 11: SEQ ID NO 2978; 2922pp; English.  
 XX  
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the

|    |      |  |      |
|----|------|--|------|
| Oy | 1140 | CCTCTCATCTCAGTGGATATATCCCTCCAAAAGGAAACAAGGCAAGGCATTATCTTTGT    | 1199 |
| Db | 1159 | CCTCTCATCTCAGTGGATATAT - CCTCCAAAAGGAAACAAGGCAAGGCATTATCTTTGT  | 1217 |
| Oy | 1200 | GTCAAAACGGGAGAGTCACCCACACTGTGTGGAAGACAAGATGCTCTCGACAGTTGCAGTAC | 1259 |
| Db | 1218 | GTCAAAACGGGAGAGTCACCCACACTGTGTGGAAGACAAGATGCTCTCGACAGTTGCAGTAC | 1277 |
| Oy | 1260 | TTACCCTTGGCTAA   | 1273 |
| Db | 1278 | TTACCCTTGGCTAA   | 1291 |

## RESULT 4

RESULTS 4  
AAT14037  
ID AAT14037 standard; cDNA; 1269 BP.

AC AAT14037;

DT 07-JUL-1996 (first entry)

DE Transcription factor-IIIa gene.

Human; transcription factor-IIIa; hTFIIIA; DNA binding protein;  
ribosome; zinc finger; rapid amplification of cDNA ends; 5'-RACE;  
primer; PCR; polymerase chain reaction; foetal brain;  
anchor primer; diagnostic; probe; transcription control;  
antitumour; cancer; therapy; ss.

OS Homo sapiens.

| XX | Key | Location/Qualifiers |
|----|-----|---------------------|
| FH |     |                     |

|    |              |            |
|----|--------------|------------|
| FT | key          | LOCATION/A |
| FT | misc_feature | 336..1115  |

11  
12  
13

\*\*\*\*\*  
tag= a

FT  
/product= zinc finger domains

polyA\_signal

W  
q -bag= b

PN EP704526-A1.

XX  
PD  
03-APR-1996.XX  
EE  
05-SEP-1995: 055P-0113908XX  
PR 05-SEP-1994. 0479-0311033XX  
PA  
(SAYAN, OTENUKA DUMRM CO T ED)

xx  
PI Fujiwara T, Ozaki K, Shimada Y, Shin S, Takeda S;

DR WPI; 1996-173033/18.  
DR P-PSDB: AAR91305.

Human Transcription Factor III A gene - useful in regulation of transcription and for diagnosis and treatment of e.g. cancer related diseases

PS Claim 2; page 10-11; 17pp: English.

The sequence encodes human transcription factor-IIIA (hTFIIIA), a DNA binding protein with 9 zinc finger domains, which is necessary for the initiation of 5S RNA gene transcription, binding to an internal control region of the 5S gene. A fuller cDNA sequence with flanking regions is given in AAT14038. A fragment lacking a 5'-portion of the gene has been isolated from a human foetal brain cDNA library (OTK7-1), and the 5'-portion of the gene has been isolated by 5'-rapid amplification of cDNA ends using primers H11-R (AAT14039), H11-E (AAT14040), H11-H (AAT14041) and AP-2 (AAT14042), complementary to anchor primer AAT14043. Reverse transcription using H-11R is followed by anchor primer ligation and PCR using AP-2 and H11-E, to give a full-length cDNA, OTK7. The gene and its encoded protein may be used in diagnosis, identification or therapy of hereditary diseases such as cancer, or other diseases resulting from

Qy 961 TCTCCAAAGCCATATCTCTCCATGAGGAAAGCGCCCTTTTGTGTGTGAACATGC 1020  
Db 960 TCTCCAAAGCCATATCTCTCCATGAGGAAAGCGCCCTTTTGTGTGTGAACATGC 1019  
Qy 1021 TGGCTGTGGCAAAACATTTGCAATGAAACAAAGTCTCACTAGGCATGCTGTTGTACATGA 1080  
Db 1020 TGGCTGTGGCAAAACATTTGCAATGAAACAAAGTCTCACTAGGCATGCTGTTGTACATGA 1079  
Qy 1081 TCCTGACAAGAAATGAAGTCAAAATCAAAATCTCGTGAAAAAC-GGAGTTTGG 1139  
Db 1080 TCCTGACAAGAAATGAAGTCAAAATCAAAATCTCGTGAAAAACGGAGTTTGG 1139  
Qy 1140 CCTCTCATCTCAGTGGATATATCCCTCCCAAGGAACAAAGGCAAGGCTTATCTTTGT 1199  
Db 1140 CCTCTCATCTCAGTGGATATAT-CCTCCCAAGGAACAAAGGCAAGGCTTATCTTTGT 1198  
Qy 1200 GTCAAAACGGAGAGTCAACCCCAACTGTGTGGAAGACAAGATGCTCTCGACAGTTGCAGTAC 1259  
Db 1199 GTCAAAACGGAGAGTCAACCCCAACTGTGTGGAAGACAAGATGCTCTCGACAGTTGCAGTAC 1258  
Qy 1260 TTACCCCTTGGC 1270  
Db 1259 TTACCCCTTGGC 1269

RESULT 5  
AAAI5407  
ID AAI5407 standard; DNA; 1213 BP.  
AC AAI5407;  
DT 04-SEP-2000 (first entry)  
XX  
DE Fragment of DNA encoding a transcription factor designated htfIIIA.  
XX  
KW Human; transcription factor; htfIIIA; DNA-binding protein;  
KW transcription; ribosomal RNA 5S gene; transcriptional control;  
XX cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200028024-A1.  
XX  
PD 18-MAY-2000.  
XX  
PF 09-NOV-1999; 99WO-FR02738.  
XX  
PR 10-NOV-1998; 98FR-0014146.  
XX  
PA (HMRI ) HOECHST MARION ROUSSEL.  
XX  
PI Bordon-Pallier F, Rocher C;  
XX  
WPI: 2000-387419/33.  
DR  
XX  
PT New nucleic acid encoding human transcription factor IIA, useful for  
PT treatment and diagnosis of cancer and inherited disease --  
XX  
PS Claim 4; Page 42; 49pp; French.  
XX  
CC The present sequence represents a fragment of a human transcription  
CC factor (designated htfIIIA) gene. The polypeptide is probably a  
CC DNA-binding protein probably involved in initiating transcription of  
CC the gene for ribosomal RNA 5S and maintaining the stability of  
CC transcription of other control genes. The htfIIIA polynucleotides and  
CC polypeptides are used to make therapeutic or diagnostic compositions  
CC for diseases associated with disorders of transcriptional control,  
CC particularly cancer or other inherited diseases. The htfIIIA  
CC polynucleotide can also be used to detect anomalies in gene  
CC transcription, particularly for diagnosis of inherited disease, also  
CC for studying diseases involving htfIIIA.

Sequence 1213 BP; 349 A; 309 C; 302 G; 253 T; 0 other;

Query Match 95.2%; Score 1211.4; DB 21; Length 1213;  
Best Local Similarity 99.9%; Pred. No. 4.5e-306;  
Matches 1212; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 61 GTGCCGGCGTCCGCGAAGGTTTCAGCAGGAGGAGCCGTTGGCCCGGCGCGGTTCCCGGC 120  
Db 1 GTGCCGGCGCGCGAAGGTTTCAGCAGGAGGAGCCGTTGGCCCGGCGCGGTTCCCGGC 60  
Qy 121 ACCTGTCTCGCACCTGGCAGCGCCCTGGCCCTGGGCTTGGAGCGCCGCGCCCTTGA 180  
Db 61 ACCTGTCTCGCACCTGGCAGCGCCCTGGCCCTGGGCTTGGAGCGCCGCGCCCTTGA 120  
Qy 181 TCCGCCGCGCGTGGTGGCGGAGTGGTCTCTTGGACCATCGCGGACGCGTTCATTGC 240  
Db 121 TCCGCCGCGCGTGGTGGCGGAGTGGTCTCTTGGACCATCGCGGACGCGTTCATTGC 180  
Qy 241 AGCCGGCAGAGTTCAGCTCCGACCCCGCCGCGCCGCTTCCAGGAGGTTTCATCTG 300  
Db 181 AGCCGGCAGAGTTCAGCTCCGACCCCGCCGCGCCGCTTCCAGGAGGTTTCATCTG 240  
Qy 301 CTCCTTCCCTGACTGCGAGCGCCCAATTACAGCAAGCCCTGGAAGCTTGACGCGCACCTGTG 360  
Db 241 CTCCTTCCCTGACTGCGAGCGCCCAATTACAGCAAGCCCTGGAAGCTTGACGCGCACCTGTG 300  
Qy 361 CAAGCACACGGGGAGAGACCATTTGTGTGTGACTATGAAGGGTGTGGCAAGGCCCTTCAT 420  
Db 301 CAAGCACACGGGGAGAGACCATTTGTGTGTGACTATGAAGGGTGTGGCAAGGCCCTTCAT 360  
Qy 421 CAGGGACTACCATCTGAGCGCGCCACATTTCTGACTCAGCAGGAGGAGGAGGAGGAGGAGG 480  
Db 361 CAGGGACTACCATCTGAGCGCGCCACATTTCTGACTCAGCAGGAGGAGGAGGAGGAGGAGG 420  
Qy 481 TGCAGCCACTGGCTGTGATCAAAATTCACACAAAATCAAACTTGAAGAAACATTTTGA 540  
Db 421 TGCAGCCACTGGCTGTGATCAAAATTCACACAAAATCAAACTTGAAGAAACATTTTGA 480  
Qy 541 ACGCAAAACATGAAATCAACAAAACAAATATATATGAGGTTTGAAGACTGTGAAGAGAC 600  
Db 481 ACGCAAAACATGAAATCAACAAAACAAATATATATGAGGTTTGAAGACTGTGAAGAGAC 540  
Qy 601 CTTTAAGAAACATCAGCAGCTGAAAATCCATCAGTGCAGCAGCATACCAATGAACCTCTATT 660  
Db 541 CTTTAAGAAACATCAGCAGCTGAAAATCCATCAGTGCAGCAGCATACCAATGAACCTCTATT 600  
Qy 661 CAAGTGTACCCAGGAGGATGTGGGAAACACTTTTGCATCACCCAGCAAGCTGAAACGACA 720  
Db 601 CAAGTGTACCCAGGAGGATGTGGGAAACACTTTTGCATCACCCAGCAAGCTGAAACGACA 660  
Qy 721 TGCCAAAGGCGCCAGGCGCTATGTATGCAAAAGGATGTTCTTGTGGCAAAACATG 780  
Db 661 TGCCAAAGGCGCCAGGCGCTATGTATGCAAAAGGATGTTCTTGTGGCAAAACATG 720  
Qy 781 GACGGAACCTTCTGAAACATGTGAGAGAAACCCATTAAGCAACACATGAAACCTCATG 840  
Db 721 GACGGAACCTTCTGAAACATGTGAGAGAAACCCATTAAGCAACACATGAAACCTCATG 780  
Qy 841 CCGGAAACATTTAAACGCAAGATTTACCTTAAGCAACACATGAAACCTCATGCCCCAGA 900  
Db 781 CCGGAAACATTTAAACGCAAGATTTACCTTAAGCAACACATGAAACCTCATGCCCCAGA 840  
Qy 901 AAGGATGTATGTGCGCTGTCCAAGAGAGGCTGTGGAAGAACCTTACTACTGTGTTAA 960  
Db 841 AAGGATGTATGTGCGCTGTCCAAGAGAGGCTGTGGAAGAACCTTACTACTGTGTTAA 900  
Qy 961 TCTCAAAGGCCATATCTCTCTCCATGAGGAAAGCCCGCTTTGTGTGTGAACATGC 1020  
Db 901 TCTCAAAGGCCATATCTCTCTCCATGAGGAAAGCCCGCTTTGTGTGTGAACATGC 960  
Qy 1021 TGGCTGTGGCAAAACATTTTCAATGAACAAAGTCTCACTAGGCATGCTGTGTACATGA 1080  
Db 961 TGGCTGTGGCAAAACATTTTCAATGAACAAAGTCTCACTAGGCATGCTGTGTACATGA 1020



